Economists Online: user requirements for a subject repository

The Network of European Economists Online (NEEO) project is a 30-month project running from September 2007 – February 2010. NEEO aims to create a central multilingual portal, Economists Online (EO), with enhanced metadata descriptions and links to the full-text research publications and primary data output of at least 500 top researchers and the economics publications of partner institutions.

In late 2007 the first user requirements study was undertaken to identify the needs of economics researchers regarding access to full-text documents and datasets; dissemination of research outputs; services to be implemented in EO; and any multilingual issues. A follow-up mid-project questionnaire was conducted in early 2009 in order to obtain early feedback from users of EO and identify any necessary refinements. A third evaluation questionnaire will be conducted near the end of the project. This article presents the results of the first two of these studies.

MICHELLE BLAKE
Liaison & Service Development Manager
London School of Economics & Political Science

Introduction

The Network of European Economists Online (NEEO) project is a 30-month project running from September 2007 – February 2010. NEEO is the flagship project of the Nereus Consortium and aims to create a central multilingual portal, Economists Online (EO).[1]

EO is based on the content of the institutional repository partners, who are:

Tilburg University (The Netherlands) – Coordinator
London School of Economics and Political Science (UK)
Erasmus University Rotterdam (The Netherlands)
German National Library of Economics/Kiel Institute for the World Economy (Germany)
Sciences Po (France)
Université Libre de Bruxelles (Belgium)
University College Dublin (Ireland)
University College London (UK)
University of Oxford (UK)
University of Warwick (UK)
Katholieke Universiteit Leuven (Belgium)
Maastricht University (The Netherlands)
Universidad Carlos III de Madrid (Spain)
Charles University in Prague, CERGE (Czech Republic)
Université Toulouse 1 Sciences Sociales (France)
Université Paris Dauphine (France)
Universität Konstanz (Germany)
Université catholique de Louvain (Belgium)
Université de Genève (Switzerland)
Columbia University (US)
European University Institute (International).

It takes an international, subject-oriented approach which aims to set standards, guarantee the quality of the information and, as a result, provide a framework that can act as a model for others to follow. The content from the partner institutions is complemented by other open access economics content including RePEc (Research Papers in Economics). The portal aims to provide an appropriate, reliable and sustainable service to various user communities, including but not limited to researchers, students and policy-makers in the public and private sectors. One of the central objectives of the project was to systematically identify the needs of these users. The project consists of three user studies. The first aimed to address the following:

1. Current researchers' needs for access to full-text documents and statistical datasets.
2. Current researchers' needs for storing and disseminating the full-text documents and statistical datasets they have produced.

3. Services to be implemented in EO.
4. Relevant multilingual and cultural issues.

The second user requirements study was undertaken halfway through the project after the initial version of the portal was running so that feedback could be obtained to allow for early refinements to be made. A third evaluation study will be conducted near the end of the project to consider the impact and success of the portal.

This article presents the results of the first two of these studies, which were carried out between September 2007 – January 2008 and March – May 2009 respectively.

Respondents for both the initial user requirements and for the follow-up mid-project study were very similar overall (see Table 1). Most respondents were academic staff (68% in 2008; 41% in 2009) and between 25 and 34 years old (35% in 2008; 37% in 2009). Respondents from the UK were quite prominent in both studies (13.3% in 2008; 18.4% in 2009) although the majority of respondents came from Spain in 2008 (25%), and the Czech Republic in 2009 (22%). In the 2009 study a total of 73% of respondents were either currently producing or expected to produce research outputs. This is slightly lower than the first study in 2008, where 91% of respondents were either producing or were likely to produce such outputs. Only 6% of respondents indicated they were unlikely to produce research outputs in the future in 2008 compared with 20% of respondents in 2009, however, this figure is likely to be associated with the fact that 21% of respondents were librarians or information professionals in 2009.

The majority of respondents in the second study indicated that they produce research outputs in English (91%). However, publication in other languages was common: 13% of respondents published in French, 12% in German, 8% in Spanish and 11% in other languages (including Dutch, Russian, Romanian and Italian). These

Academic Role:	**2008 %**	**2009 %**
Academic staff (faculty member with research and teaching responsibilities)	68.1%	40.8%
Research staff (research-only)	10.5%	21.1%
Research student (doctoral level student)	12.6%	21.1%
Student (masters or undergraduate level)	3.8%	6.6%
Librarian or information professional	3.2%	6.6%
Other	1.9%	3.9%
Range of age:	**2008 %**	**2009 %**
under 25	4.7%	13.2%
25-34	34.9%	36.8%
35-44	30.6%	22.4%
45-54	15.8%	19.7%
55 or over	14.1%	7.9%
Country:	**2008 %**	**2009 %**
Spain	25.1%	3.9%
UK	13.3%	18.4%
Belgium	9.8%	9.2%
Germany	9.0%	17.1%
Netherlands	8.4%	7.9%
France	6.2%	5.3%
Czech Republic	4.9%	22.4%
Italy	3.8%	9.2%
US	3.4%	1.3%
Ireland	2.8%	1.3%
Norway	1.5%	0.0%
Others	11.8%	3.9%
Production of research in Economics:	**2008 %**	**2009 %**
I currently produce research outputs.	82.7%	55.3%
I expect to produce research outputs in the near future.	9.2%	18.4%
I have produced research outputs in the past but currently inactive.	2.1%	6.6%
No, I am unlikely to produce research outputs.	6.0%	19.7%

Table 1. Sample characteristics from 2008 and 2009 results

findings are similar to those of the first study in 2008. This illustrates the diversity of the audience whose needs EO is designed to meet.

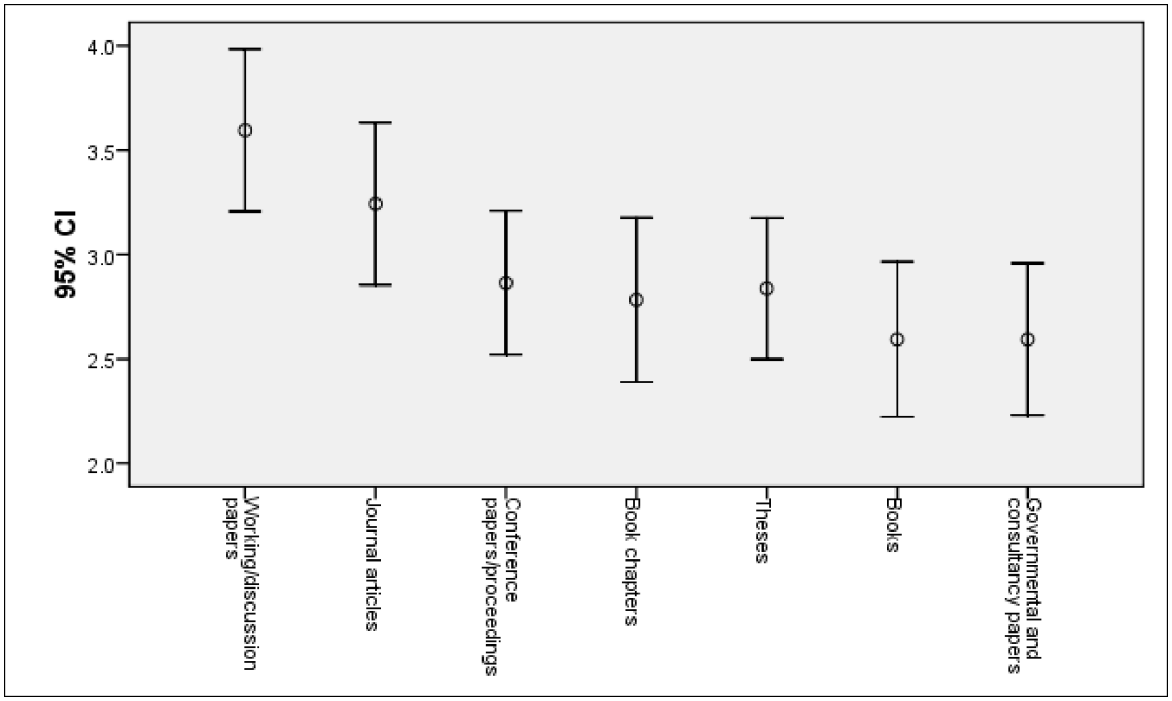
Content

Respondents to the online survey (see Figure 1) and participants of focus groups agreed that the research material planned for inclusion in EO was very comprehensive, comprising journal articles, books, book chapters, working papers, discussion papers, theses, conference proceedings and datasets. In fact, 91.6% of the online survey respondents stated there was no need to include other types of research output in the service. However, when users were asked in the second study if the portal provided a good selection of content across the various types of material it aimed to include, the results were less promising.

When comparing these findings with that of the first study in 2008, 26% of respondents found working papers hard to access and 21% found discussion papers hard to access. Respondents in 2009 valued the portal most highly in these areas. While journal articles were considered the least difficult to access online in 2008 (only 2.6% of respondents reported difficulties), EO is also considered relatively strong in providing this material.

The first study found that book chapters and books were considered most difficult to access online (with 67% and 66% of respondents respectively considering such materials difficult to locate). As a result, EO appears to have not yet exploited its potential added value in this area. The interest in books and book chapters reflects the limited availability of economics books online; for example, only 0.2% of the entire RePEc database corresponds to downloadable books[2]. However, most interviewees were doubtful whether free access to such material could be achievable due to copyright concerns and commercial interests from both authors and publishers.

Conference proceedings, discussion and working papers were of interest to the interviewees in as much as they usually contain ongoing research and could therefore provide an overview of current topical and methodological trends within the field of economics, as well as of the research process followed by some well-known authors. Of the three of them, conference proceedings appeared to be the most easily accessible online; however, it was mentioned that although conference organizers usually put the presentations online, those papers do not tend to be complete – they are



NOTE: N=49 to 61. Scale represents whether the user thinks that Economists Online provides a good selection of content, ranging from 1=Not at all and 5=Completely. Responses are ranked by average score in descending order.

Figure 1. User rating – content quality by source

mostly limited to abstracts or summaries – nor are they accessible for long periods of time. In addition, although discussion and working papers are more easily accessible through institutional and personal websites, some researchers were interested in accessing other kinds of 'working papers', understood as specialized technical reports produced for governmental and private institutions (usually evaluation and consulting reports), which are currently difficult to access online.

Although theses were reported as one of the most difficult outputs to access online, they were not as highly valued as other research outputs.

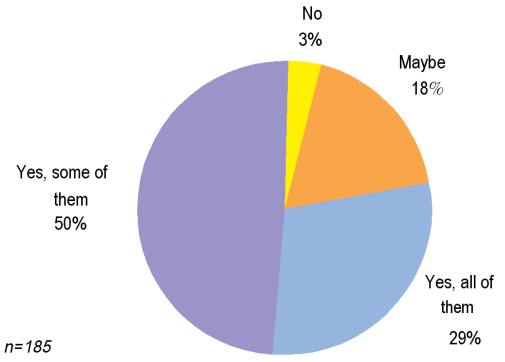


Figure 2. Disposition to submit produced datasets into a repository

Datasets

The first study asked respondents about access to datasets. Respondents indicated that if the service could not only link articles to their respective datasets but also offer a datasets portal, then it would be a key feature of Economists Online.

However, when dataset creators were asked if they would be willing to submit their datasets into an institutional repository, Figure 2 shows that there was a reluctance to do so (only 29% indicated they would be willing to submit all their datasets). When respondents were asked to state under what conditions they would be willing to submit their datasets, 75% wanted a clear statement of dataset creators' rights. They also desired a mechanism to control how end-users would access datasets in EO, either by stating clearly the permissions in place to use the data (64.9%) or by making registration compulsory (36.8%). In contrast, only a minority (23.7%) of respondents were happy to make all datasets available.

Figure 3 shows that dataset creators appear to have two main concerns in relation to submitting their work to EO: the risk of not being credited for their work, and losing the opportunity of producing more results based on their data. Even if a researcher had already produced an article, giving away his/her data might allow other researchers to produce a similar piece of work more quickly than the dataset creator.

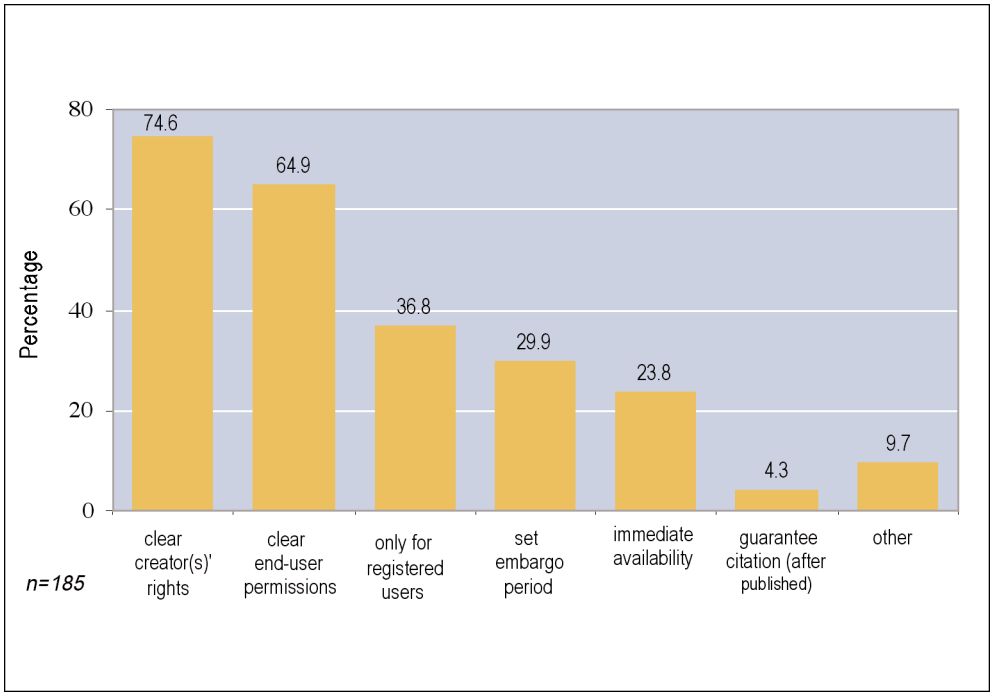


Figure 3. Conditions demanded before submitting datasets

In addition to losing the opportunity of producing more research outputs, participants also feared that free access to datasets might facilitate a widespread use of their work without any recognition. The central role datasets play in empirical research means they were not seen solely as an input for publications but also as a product in themselves, whose authorship needs to be credited. The general consensus, then, was that researchers would be willing to include datasets in EO only if creators are guaranteed proper acknowledgement.

The most common mechanism by which participants thought the above concerns could be somewhat minimized was to wait until their own research was released, preferably through an article in a peer-reviewed journal.

Datasets for the project will be linked from articles in EO and will be stored in a DataVerse repository[3].

Services

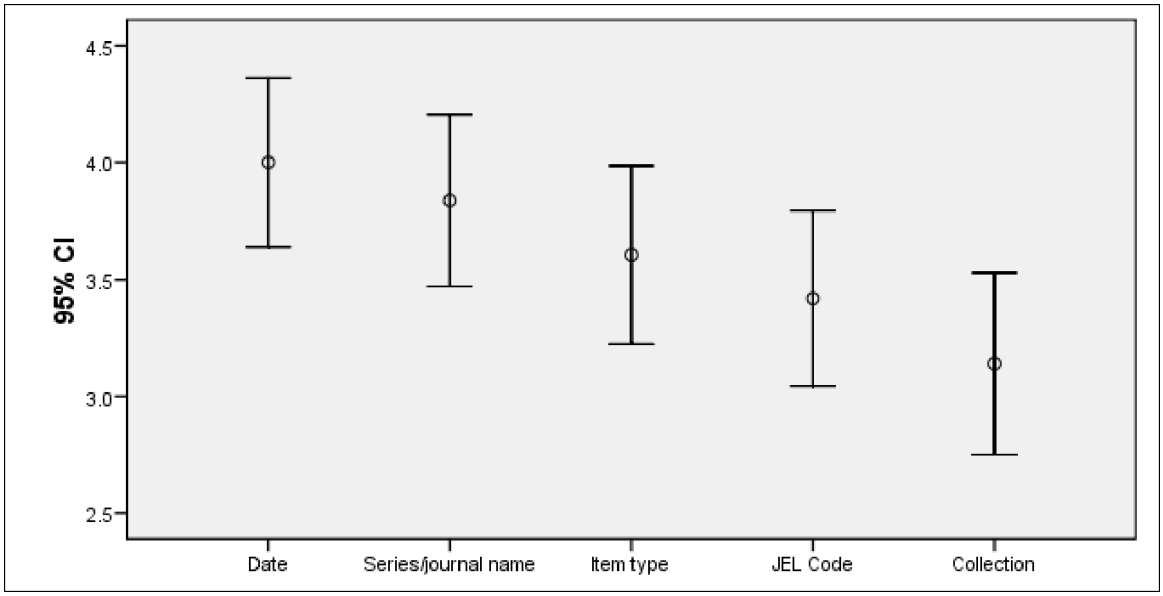
During the second study, users were asked about a range of services which had been implemented into the first version of EO. It should be noted that, since the user study was completed, substantial changes have been made to the interface of EO, some of which were based on feedback received from this study.

In the 2008 study (see Figure 4), respondents reported that searches by author, title, full text, abstract and subject were the most valuable search options. These requirements were confirmed in the second user study of current and proposed facet options. Initially, a search by date and *Journal of Economic Literature* (JEL) code were considered less valuable. However, the findings suggest that while users continue to favour the JEL search option less, they actually do value searches by date in the current version of the gateway.

The 2008 study found that respondents most favoured links to full-text for freely available titles (see Figure 5). This service also proved most popular in the survey results from 2009. Similarly, in 2008 users favoured links to fee-based publications less, a service that has also proven less popular in the first version of the portal.

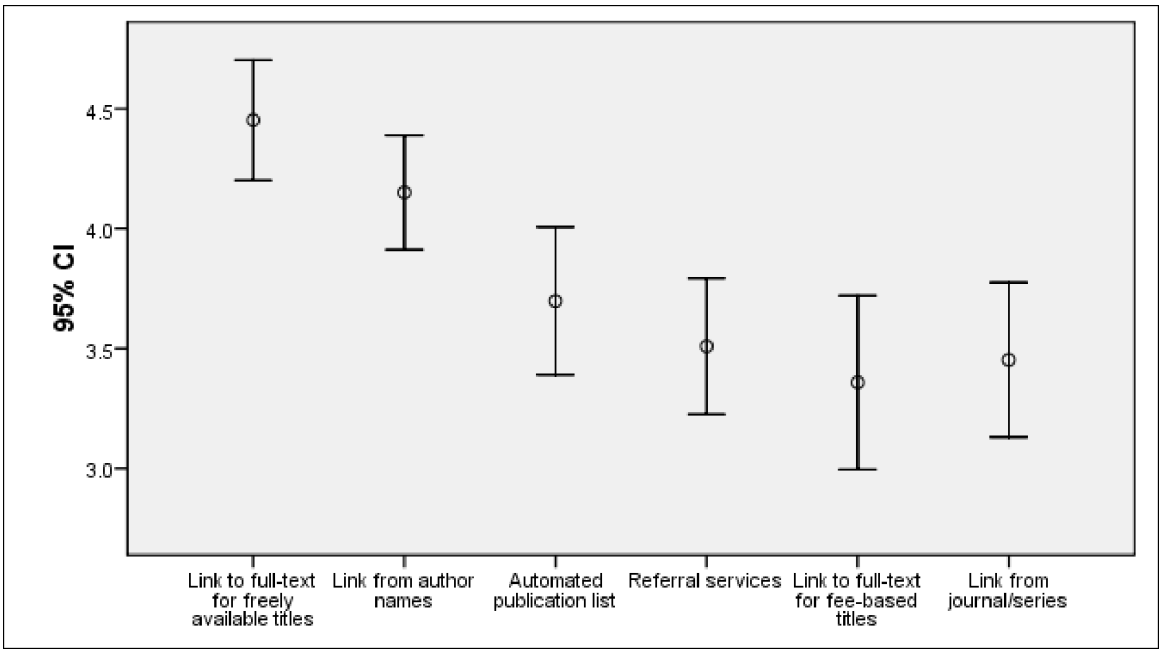
Metrics

Only a minority of economists surveyed in 2008 reported having access to information on how often their work is cited or used (37%). Existing or likely contributors to EO were asked what kinds of metrics they would find useful in relation to their publications. The most popular metric was total number of downloads, followed by downloads by month/year and the number of unique individuals accessing work. Less popular was the country of



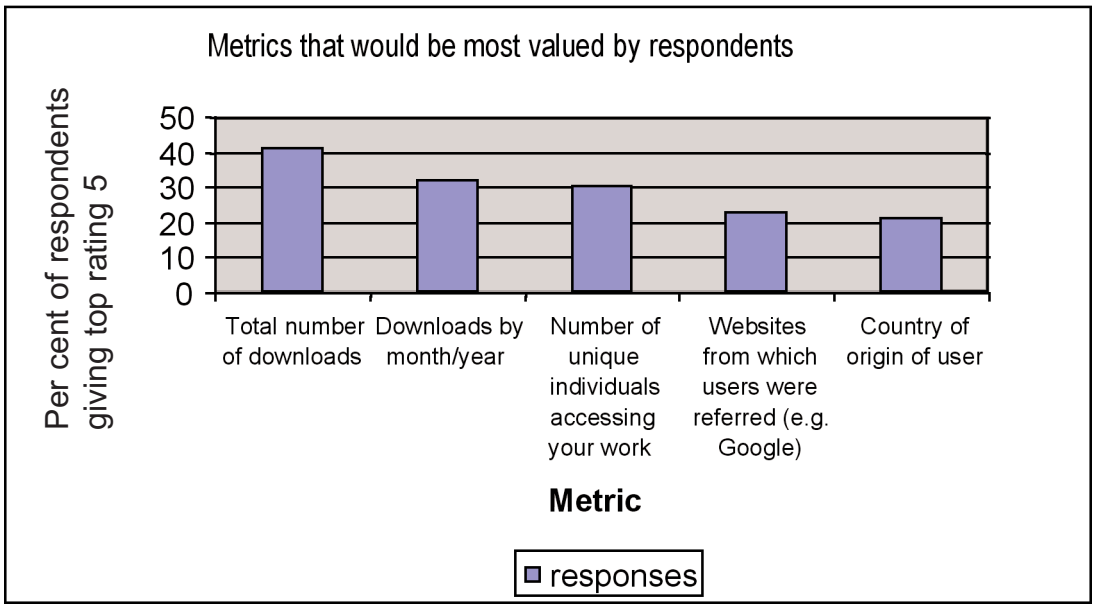
NOTE: N=76. Scale represents the user's opinion on how useful the facet options were for expanding or limiting their search, from 1=Not at all to 5=Completely. Responses are ranked by average score in descending order.

Figure 4. User ratings – facet options



NOTE: N=76. Scale represents the user's opinion on the usefulness of the service, ranging from 1=Not at all to 5=Completely. Responses are ranked by average score in descending order.

Figure 5. User ratings – services provided by Economists Online



NOTE: N=76. The top rating 5 indicates that the user considers this metric 'completely' important, rather than rating 1, 'Not at all' important.

Figure 6. Metrics favoured by survey respondents

origin of users accessing work and websites from which users were referred. Figure 6 illustrates the proportion of respondents who gave the ranking of 5 to each metric, indicating that it is of high importance to them.

Dissemination

Most economics researchers who participated in the 2008 survey demonstrated that they make use of open access sites such as institutional and subject repositories as well as personal websites.

The majority of researchers (70%) indicated that most of their produced material is already available through those sources. However, only a minority (30%) estimated that all their outputs were accessible in this manner. Discussion papers and working papers are commonly freely accessible online, while journal articles are not traditionally open access. The survey questions did not ask respondents to differentiate between types of research output they produced or made open access. However, analysis of the pilot EO service metadata records conducted for the JISC-funded VERSIONS Project suggests that 23% of EO authors' research outputs were discussion papers and working papers[4].

Focus group participants expressed the view that one of the main reasons for this patchy use of open access services was copyright. Although participants said they do make the great majority of their discussion and working papers available to anyone interested, they expressed serious concerns about their right to provide final drafts of articles published in peer-reviewed journals or books because of potentially breaching any previous agreement with publishers.

Respondents were asked in the 2008 study to indicate the search services they use when conducting research. Figure 7 shows that the most popular search services were Google and Google Scholar, followed by RePEc and the Social Science Research Network (SSRN). Therefore, it was not unexpected that when researchers were asked to state through which online services they would like to see their work disseminated, the majority preferred Google and Google Scholar alongside more specialized online bibliographic databases such as RePEc and SSRN. Library services such as catalogues, databases or repositories were of less importance for respondents.

Three questions sought user feedback on the value of the EO portal for the dissemination of their work. Firstly, respondents were asked if they had already contributed, or were planning to contribute, to EO, and if they believed the service would increase the visibility of their work.

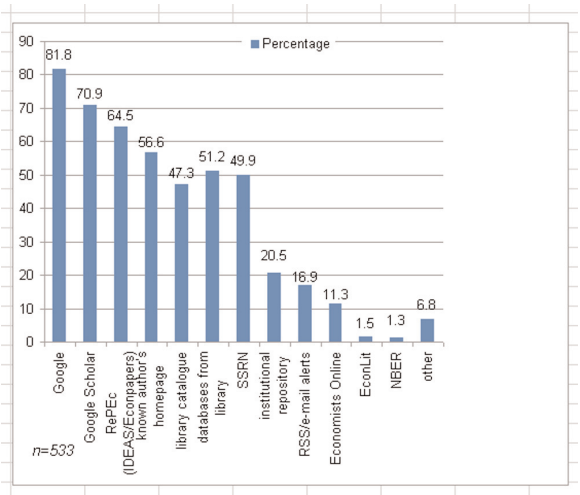
Of the individuals who were not yet contributors, 88% believed that it would increase the visibility of their work. This provides strong support in favour of the portal.

Secondly, users were asked, if they had not yet contributed to EO, would they be willing to submit their research to this open access service? 89% indicated that they would be willing to do this. Again, this shows strong support among the target community for the EO service.

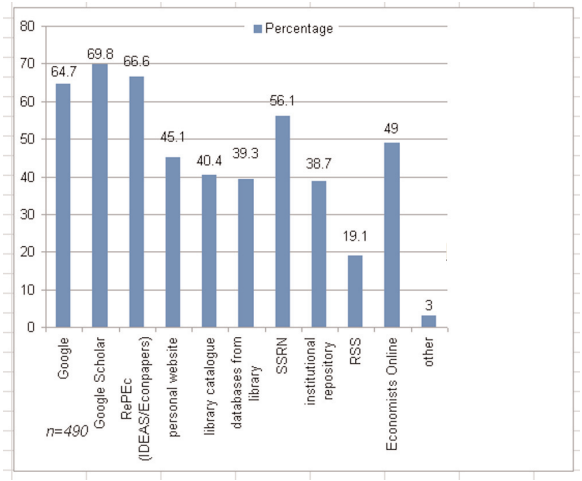
Finally, the survey asked whether participants would recommend participating in EO to a colleague. Most respondents (73%) indicated that they would do so – indicating widespread positive support for the portal.

Conclusion

The two user studies that have been completed have been able to inform the development of the EO portal. The portal that was used for the 2009 survey was a



Services used to discover economists' research outputs



Researchers' preferred service to disseminate their research outputs

Figure 7. Services used by researchers and researchers preferred services for dissemination

soft launch version and since its launch many technical difficulties and issues with design have already been resolved. In addition, it is expected that planned future developments of the portal will continue to improve user satisfaction.

The third and final user study is due to be completed by the end of the project in February 2010. This study will follow on from the mid-project 2009 study and gain feedback on the value of EO in all areas, including datasets.

References

1. Economists Online (EO) http://nereus.uvt.nl/eo (accessed 28 September 2009).
2. RePEc – IDEAS, 2008: http://ideas.repec.org/ (accessed 5 October 2009).
3. DataVerse repository: http://thedata.org/ (accessed 28 September 2009).
4. VERSIONS Project. Publications list analysis report, November 2007: http://www.lse.ac.uk/versions (accessed 28 September 2009).

■ **Michelle Blake**
Liaison & Service Development Manager
London School of Economics & Political Science
LSE Library
10 Portugal Street
London WC2A 2HD, UK
E-mail: M.Blake@lse.ac.uk